Common genetic variants associated with pancreatic adenocarcinoma may also modify risk of pancreatic neuroendocrine neoplasms

Ofure Obazee [1], Gabriele Capurso [2], Francesca Tavano [3], Livia Archibugi [2], Antonio De Bonis [4], William Greenhalf [5], Tim Key [6], Claudio Pasquali [7], Anna Caterina Milanetto [7], Thilo Hackert [8], Paola Fogar [9], Valbona Liço [7], Christos Dervenis [10], Rita T. Lawlor [11], Luca Landoni [12], Maria Gazouli [13], Carlo Federico Zambon [14], Niccola Funel [15], Oliver Strobel [8], Krzysztof Jamroziak [16], Cinzia Cantù [11], Ewa Małecka-Panas [17], Stefano Landi [18], John P Neoptolemos [5], Daniela Basso [9], Renata Talar-Wojnarowska [17], Maria Rinzivillo [2], Angelo Andriulli [3], Federico Canzian [1]*, Daniele Campa [18]*

Affiliations:

[1] Genomic Epidemiology Group, German Cancer Research Centre (DKFZ), Heidelberg, Germany

[2] Digestive and Liver Disease Unit, S. Andrea Hospital, 'Sapienza' University of Rome, Rome, Italy

[3] Division of Gastroenterology and Research Laboratory, San Giovanni Rotondo, Italy

[4] Department of Surgery, IRCCS Scientific Institute and Regional General Hospital "Casa Sollievo della Sofferenza", San Giovanni Rotondo, Italy

[5] Department of Molecular and Clinical Cancer Medicine, Institute of Translational Medicine, University of Liverpool, United Kingdom

[6] Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[7] Pancreatic and Digestive Endocrine Surgery - Department of Surgery, Oncology and Gastroenterology -DiSCOG, University of Padova, Italy

[8] Klinik für Allgemein-, Viszeral- und Transplantationschirurgie, Im Neuenheimer Feld 110, D-69120 Heidelberg, Germany

[9] Department of Laboratory Medicine, University-Hospital of Padova, Italy

[10] Department of Surgical Oncology and Hepatobiliary Surgery, Metropolitan General Hospital, Pireas, Greece

[11] ARC-NET Center for Applied Research on Cancer, University and Hospital Trust of Verona, Verona, Italy

[12] Department of Surgery, Pancreas Institute, University and Hospital Trust of Verona, Verona, Italy

[13] Department of Basic Medical Sciences, Laboratory of Biology, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece

[14] Department of Medicine - DIMED, University of Padova, Italy

[15] Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Italy

[16] Department of Hematology, Medical University of Lodz, Poland

[17] Department of Digestive Tract Diseases, Medical University of Lodz, Poland

[18] Department of Biology, University of Pisa, Italy

* Authors share co-senior authorship.

Running title: PDAC GWAS susceptibility loci may modify pNEN risk

Keywords: pNEN, PANDoRA consortium, PDAC, genetic susceptibility

Financial support

This study was supported by intramural funds from the German Cancer Research Centre (DKFZ, Heidelberg, Germany) awarded to F.C. G.C. was supported by AIRC grant IG 2015, Id 17177.

Corresponding author

Federico Canzian, Genomic Epidemiology group, German Cancer Research Centre (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. Phone: +49-6221-421791. Fax: +49-6221-421810. E-mail: f.canzian@dkfz.de

Disclosure of potential conflicts of interest

The authors declare that there is no conflict of interest (financial, professional or personal) that could be perceived as prejudicing the impartiality of the research reported.

Abstract

Pancreatic neuroendocrine neoplasms (pNEN) account for less than 5% of all pancreatic neoplasms and genetic association studies on susceptibility to the disease are limited. We sought to identify possible overlap of genetic susceptibility loci between pancreatic ductal adenocarcinoma (PDAC) and pNEN; therefore, PDAC susceptibility variants (n=23) from Caucasian genome-wide association studies (GWAS) were genotyped in 369 pNEN cases and 3,277 controls from the PANcreatic Disease ReseArch (PANDoRA) consortium to evaluate the odds associated with pNEN risk, disease onset and tumor characteristics. Main effect analyses showed four PDAC susceptibility variants – rs9854771, rs1561927, rs9543325 and rs10919791 to be associated with pNEN risk. Subsequently, only associations with rs9543325, rs10919791 and rs1561927 were noteworthy with false positive report probability (FPRP) tests. Stratified analyses considering age at onset (50 year threshold), showed rs2736098, rs16986825 and rs9854771 to be associated with risk of developing pNEN at a younger age. Stratified analyses also showed some SNPs to be associated with different degrees of tumor grade, metastatic potential and functionality. Our results identify known GWAS PDAC susceptibility loci, which may also be involved in sporadic pNEN etiology and suggest that some genetic mechanisms governing pathogenesis of these two entities may be similar, with few of these loci being more influential in younger cases or tumor subtypes.

Summary: This study identifies an overlap of susceptibility loci between PDAC and pNEN which may provide insights on potentially useful markers for risk stratification and tumor characterization among healthy individuals and pNEN patients respectively.

Introduction

Pancreatic neuroendocrine neoplasms (pNEN) arise from islet cells and comprise less than 5% of all new pancreatic neoplasms (1). Heterogeneous in tumor behavior and clinical symptoms, pNENs account for 10% of neuroendocrine tumors and have a 5-year mortality rate of 60% (1,2). Compared to the more common exocrine pancreatic ductal adenocarcinoma (PDAC), pNENs are characterized by a rather 'silent' clinical course and generally present at a relatively earlier age (median age at diagnosis 53 to 60 years, and 72 years for PDAC) (1). Majority (40-90%) of pNENs are non-functional (i.e. do not secrete hormones that cause systemic effects) thus constituting a clinical and prognostic challenge for physicians. Surgical resection is currently the primary curative therapy option for both malignancies, however, most patients unfortunately exhibit unresectable tumors at diagnosis (3), stressing the need for improved risk stratification and timely diagnostic biomarkers. Familial clustering observed in population-based studies among PDAC and other cancer patients provide evidence of an inherited basis for sporadic pNEN (1). Although there is no clear link between environmental exposures and risk of developing pNEN, an overlap of risk factors in PDAC and pNENs is likely (4). Increased interest in these neoplasms has become evident in the last decade, but despite this, molecular understanding of pNENs is still insufficient to drive clinical interventions. Recent exome and whole genome sequencing studies describing the genetic basis of pNENs among Caucasians (5,6) have implicated somatic mutations in four core pathways: activation of mTOR signaling, DNA damage repair, chromatin modification and altered telomere length in tumorigenesis. Genome wide association (GWAS) or candidate gene or pathway studies investigating the

etiology of pNENs have been reasonably limited by small sample size, when compared to PDAC for which 23 common risk variants have so far been identified through GWAS or studies focusing on genes situated in well-known pleiotropic regions (7–11). Based on the rationale that known common genetic susceptibility loci for sporadic PDAC may also modify the risk of developing pNEN, we genotyped 369 pNEN cases from a retrospectively enrolled population within the PANDoRA consortium and compared genotype frequencies of all known PDAC GWAS susceptibility variants to over 3,200 control subjects from the same consortium. We tested their possible association with overall sporadic pNEN risk and within pNEN subgroups.

Materials and methods

Selection of subjects and polymorphisms: Demographic characteristics subjects within the PANDoRA consortium have been described previously (12) although the number of neuroendocrine cases within PANDoRA has increased since that publication. Written informed consent and biospecimens (blood, tissue or genomic DNA) were obtained for all subjects and ethical approval for the PANDoRA study protocol received from the Ethics commission of the Medical Faculty Heidelberg. Only sporadic (non-familial) pNEN cases not associated with genetic syndromes such as multiple endocrine neoplasia type 1 (MEN-1), von Hippel-Lindau syndrome (VHL), von Recklinghausen disease (neurofibromatosis NF-1), and tuberous sclerosis complex (TSC) were used in this study. Twenty-three (23) single nucleotide polymorphisms (SNPs) reaching genome-wide significance and suggestive levels (up to 5×10^{-7}) were selected from GWAS publications in PDAC among Caucasians (7–10). pNEN tumors were classified according to the World Health Organization 2010 classification system (13).

Genotyping and quality control: SNPs were genotyped in 369 pNEN cases and 3,277 controls using TaqMan allelic discrimination (Applied Biosystems) and KASPar (LGC Genomics) assays according to manufacturers' instructions. Post-PCR allelic discrimination was done on the 7900HT Real-Time PCR system (Life Technologies) and data analyzed using the affiliated SDS software (v.2.3). Internal replicates (7% of all samples) and negative controls were included to assess the fidelity of genotype calls. Deviation from Hardy Weinberg equilibrium (HWE) was checked using the Pearson's χ^2 test.

Statistical analyses: We evaluated associations (odds ratios (OR) and 95% confidence intervals (CI)) between SNPs and pNEN risk using unconditional multivariate logistic regression in STATA v.11 (StataCorp LP). Tests were done using a codominant inheritance model, with adjustment for potential confounding variables - age, geographical origin and gender. We used a nominal significance threshold ($p<0.05$) since the selected SNPs were based on prior evidence of their association with pancreatic cancer. Noteworthiness of SNP associations with $p<0.05$ were tested using the Bayesian false positive report probability (FPRP) test which takes the observed *P*-value, statistical power of the test, and prior probabilities for the associations into account (14). All selected SNPs bore a high prior possibility of being associated with pNEN risk ($\pi = 0.2$) as recommended by Wacholder and colleagues (14), and only SNP associations with $p<0.05$ and $\pi<0.2$ were considered 'noteworthy'. To evaluate the effect of SNPs on pNEN onset, we first performed a case-only analysis comparing genotype distributions among cases (>50 years) with their younger counterparts (≤50 years) while adjusting for gender and geographical origin. We also compared genetic associations from three additional models: cases ≤50 years versus controls of all ages (EOP1), cases >50 years versus controls of all ages (TOP), and cases up to 50 years versus controls up to 50 years (EOP2). Associations between SNPs and tumor grade and stage were estimated by case-only analyses. For tumor stage, we compared stage IV tumors versus all other tumors (stage I, IIa, IIb, IIIa, IIIb) based on European Neuroendocrine Tumor Society (ENETS) classification system (15) and for grade, we compared well differentiated tumors (grade 1 and 2) versus poorly differentiated tumors (grade 3) based on the World Health Organization (WHO)/ENETS guidelines (16). Tumor functional status (e.g. insulinoma,

gastrinomas, etc) was available for 54% of all cases in this study, out of which 62% were non-functioning. We compared genotype frequencies of all SNPs in the latter with functioning tumors.

Microarray data screening: To gain additional insight to the transcriptomic profiles of these gene regions in human pancreatic islet cells, we explored expression profiling data from the Gene Expression Omnibus (GEO) database (Accession number GSE43795); this study characterized mRNA and microRNA expression profiles in solid-pseudopapillary neoplasms of pancreas (n=14), ductal adenocarcinoma (n=6), pancreatic neuroendocrine tumors (n=6) and non-neoplastic pancreatic tissue samples (n=5) (17). Results were downloaded in text format and GEO2R queries (http://www.ncbi.nlm.nih.gov/geo/geo2r/) applied to assess differentially-expressed genes in pNEN compared to non-neoplastic pancreas tissue samples. GEO2R is an R programming language-based dataset analysis tool that compares two or more groups of samples under the same experimental conditions and analyzes almost any GEO series (18). *P* values were adjusted to correct for false-positives due to multiple testing using Benjamini and Hochberg false discovery rate method and genes that met the cut-off criteria of $P_{adjusted} < 0.05$ and $|logFC| > 1.0$ were screened out as differentially-expressed genes.

Results

Baseline characteristics of the study population are summarized in table I. The median age at diagnosis in this study was 57 years, congruent with previous literature reports. Genotypes for most SNPs, except rs505922 (P = 0.04) were distributed in accordance to Hardy-Weinberg equilibrium (HWE) among controls. Duplicated samples showed genotyping concordance rate of 99.6%. Genotyping data for those subjects with genotyping completion rates below 75% (n = 35) were excluded from all statistical analyses.

Overall pNEN risk analyses*:* Minor allele frequencies, genotype distributions, risk estimates (odds ratios (OR) and 95% confidence intervals (CI)) of all polymorphisms investigated in this study are summarized in table II. rs9543325 was associated with higher risk of pNEN (OR 1.63, 95% CI (1.10-2.47), P_{hom} = 0.02). rs10919791, rs1561927 and rs9854771 were associated with lower risk of pNEN (OR 0.65, 95% CI (0.46-0.93), P_{het} = 0.02; OR 0.71, 95% CI (0.55-0.92), P_{het} = 0.01 and OR 0.76, 95% CI (0.58-0.99), P_{het} = 0.04 respectively). Subsequent testing of these findings using false positive report probability (FPRP) with a prior of 0.2 showed that only associations with rs10919791 (FPRP = 0.13), rs9543325 (FPRP = 0.15) and rs1561927 (FPRP = 0.07) were noteworthy. The association with rs9854771 failed the noteworthiness test (FPRP = 0.25).

Stratified analyses based on age at pNEN onset: With stratification according to age at disease onset (≤50 years vs. >50 years) among successfully genotyped pNEN cases, we observed

significant differences between genotype groups for rs16986825 (*ZNRF*) (OR 2.67, 95% CI (1.57-4.55), P_{het} = 0.0003 and OR 2.27, 95% CI (1.37-3.77), *P* = 0.002 in the recessive model); rs2736098 (*TERT*) (OR 3.71, 95% CI (1.36-10.10), P_{hom} = 0.01; and rs9854771 (*TP63*) (OR 1.85, 95% CI (1.11-3.10), P_{het} = 0.02 and OR 1.69 , 95% CI (1.04-2.75), *P* = 0.03 in the recessive model (table III). FPRP tests were similarly applied using a prior probability of 0.2 and associations between early pNEN onset and rs16986825 (FPRP < 0.01 in both models), rs2736098 (FPRP = 0.08), and rs9854771 (FPRP = 0.14), were noteworthy. For rs9854771, our observed association for the recessive model was not noteworthy (FPRP = 0.22). Since pNENs are diagnosed roughly 10 years earlier than PDAC, we also performed the same stratified analyses with threshold at 40 years, and only the *TERT* SNPs (rs2736098 and rs2853677) were significantly associated with higher risk of developing pNEN at 40 years or younger, albeit with extremely small group sizes (results not shown). Results from alternative analyses of disease onset using three additional models: cases ≤50 years versus controls of all ages (EOP1), cases >50 years versus controls of all ages (TOP), and cases up to 50 years versus controls up to 50 years (EOP2) are shown in supplementary table I.

Stratified analyses based on tumor characteristics (grade, stage and functional status): With stratification based on disease grade, rs10919791 and rs16986825 were associated with well differentiated pNEN tumors (OR 5.11, 95% CI (1.13 - 23.14), *P* = 0.03 and OR 7.44, 95% CI (1.09 - 50.5), *P* = 0.04 respectively). rs1561927, rs9854771 and rs351365 were also associated with tumors with nodal metastasis (i.e. advanced stage) (OR 2.39, 95% CI (1.18 - 4.85), *P* = 0.02; OR 0.46, 95% CI (0.24 - 0.88), *P* = 0.02; and OR 4.62, 95% CI (1.16 - 18.38), *P* = 0.03) respectively.

However, only the association between rs9854771 and stage appeared reliable, considering that the size of genotype groups was much smaller for other SNPs. FPRP tests applied using a prior probability of 0.2 showed the association between tumor grade and rs10919791 (FPRP = 0.19) to be noteworthy. For rs16986825, the association was not noteworthy (FPRP > 0.20). Regarding tumor stage, associations with rs1561927, rs9854771 and rs351365 (recessive model only) (FPRP=0.11, 0.16 and 0.19 respectively) were noteworthy. Two variants rs7310409 (*HNF1A*) and rs1517037 (*GRP*) also appeared to be associated with tumor functional status (OR 0.36, 95% CI (0.16-0.79), *P* = 0.01 and OR 0.37, 95% CI (0.16-0.85), *P* = 0.02 respectively (table III). Both associations were noteworthy at the 0.2 threshold (FPRP < 0.08). Associations between all 23 SNPs and tumor characteristics are shown in supplementary table II.

Differentially-expressed genes based on microarray data screening: Reasoning that the additional gene expression profiles of all 26 genes harboring or located close to the SNPs investigated in this study would amplify our ability to identify relevant underlying biological processes, we screened a GEO superseries dataset including 6 pNEN, 5 healthy pancreas (controls) and 6 ductal adenocarcinoma tissue samples. Based on the inclusion criteria of $P_{adjusted} < 0.05$ and $|logFC| > 1.0$, five of these genes – nuclear receptor subfamily 5 group A member 2 (*NR5A2*), v-myc avian myelocytomatosis viral oncogene homolog (*MYC*), Krüppel-like factor (KLF) 5 (*KLF5*) and chymotrypsinogen genes, *CTRB1* and *CTRB2* were under expressed in pNEN compared to normal pancreas tissue (table V). Supplementary table III shows expression data for all 23 SNPs. Our analyses are based on only a few samples, therefore caution should be applied in the interpretation of these results.

Discussion

By investigating known GWAS PDAC susceptibility loci (MAF≥5%), our study has identified the possible role of these loci in the risk of developing pNEN. Combined with gene expression analyses, we report for the first time, few loci which may also be involved in genetic mechanisms influencing pNEN tumorigenesis, age at onset and tumor characteristics. These findings confirm that some susceptibility loci are shared by both entities. Interestingly, most known environmental risk factors associated with increased pNEN risk are also associated with PDAC risk (4). Although rarity and lack of predictive biomarkers for early-stage diagnosis or risk stratification of healthy individuals in both entities are two common features of these neoplasms, understanding their molecular mechanisms remains essential and advantageous. It is common albeit controversial knowledge that pancreatic ductal cells are 'potential facultative stem cells' capable of being reprogrammed into cells that closely resemble islet cells *in vivo* (19,20). Whether the shared genetic susceptibility is due to the origin of both these tumor types from a common pluripotent precursor cell, or a phenomenon linked with the reported capacity of islet cells to transdifferentiate to ductal cells (21,22) is currently unknown. Among explanations proposed for the transdifferentiation of pancreas cells, a largely favored one in pathogenesis is based on the hypothesis that tumor stroma mimics the regulatory role of pancreatic embryonic mesenchyme in duct and endocrine development, to induce the switch between ductules and endocrine cells (23,24). A second hypothesis based on multiple lines of evidence also suggests that tumor cells may originate from a common pluripotent cell which matures into two phenotypically different cell lines (25), and this pluripotent cell expresses

transcription factors such as *PDX1* and *SOX9* which are involved in pancreatic development and homeostasis (26). Our study, which comprised of 369 pNEN cases and 3,277 healthy controls from the international PANDoRA consortium, represents the largest of such studies performed to date. Recent attempts to delineate the mutational landscape of pNENs through whole genome sequencing (6), have implicated activation of mTOR signaling as one of the core pathways commonly altered in these neoplasms.

In overall risk analyses, four SNPs (rs9854771, rs9543325, rs1561927 and rs10919791) showed associations with sporadic pNEN at $p<0.05$. Adjustment for multiple comparisons was not applied to the nominal significance threshold ($p<0.05$) since selected SNPs were based on prior evidence of their association with pancreatic cancer. Using the FPRP model in overall analyses, only associations with rs10919791, rs9543325 and rs1561927 reached noteworthiness with a prior probability of 0.2.

rs1561927 in *MIR1208*/*PVT1* was associated with a 29% lower risk of developing pNEN among heterozygotes, which is similar to the original report in PDACs (allelic OR 0.87 95% CI (0.83–0.92) (9)). This intronic SNP maps to a locus associated with suggested long-distance interactions with *MYC* and *PVT1* promoters in multiple cancers (9). rs9543325, an intergenic SNP in chr13q22.1 which is frequently deleted in several cancers (8) was associated with a 37% increased risk of developing pNEN (for subjects with CC genotype), higher than risk estimates from the original study in PDACs (OR 1.26 95% CI (1.18–1.35) (8). Genes closest to this locus are *KLF5, KLF12, PIBF1, DIS3* and *BORA,* which range up to 586kb in distance from rs9543325. Krüppel-like factor (KLF) 5 and 12 are members of a family of zinc-finger transcription factors reported to exhibit tumor-suppressor and oncogenic activity respectively in various human

cancers (27). KLF5 is involved in transcriptional activation of PI3K/Akt signaling, *BORA* (encoding aurora kinase A activator) is known to regulate cell proliferation and overexpressed in tumors (28). *PIBF1* (encoding progesterone immunomodulatory binding factor 1) is speculated to be involved in progesterone-dependent immunomodulation with higher levels in tumors (28). Aberrant expression of *DIS3* has been implicated in many cancers and plays a crucial role in gene regulation and small RNA processing (29). rs10919791 lies in the first intron of *NR5A2* (chr1q32.1) which plays crucial roles in pancreatic function and development (30). In the current study, rs10919791 was associated with a 35% lower risk of developing pNEN among heterozygotes, similar to the original estimate among PDACs (OR_{het} 0.76 (0.68–0.84)) (8). The bioinformatics approach with a GEO2R microarray dataset including pancreatic neuroendocrine tumors (n=6) and non-neoplastic pancreatic tissue samples (n=5), was used to augment results from our study and explore the plausible role of these genes in pNEN development. Interestingly, *MYC*, *KLF5*, *NR5A2*, *CTRB1* and *CTRB2* appeared down-regulated among pNENs compared to non-neoplastic pancreatic tissue in the microarray dataset, further implicating these genes in pNEN tumorigenesis. To substantiate these findings, we performed a similar comparison of expression profiles of these genes in PDAC (n=6) versus healthy pancreas, which showed differential expression of all genes but *KLF5* in the same manner. To the best of our knowledge this is the only publicly available dataset that includes expression data of pNEN. The data we observed in this dataset corroborate our findings of the association analysis. Nevertheless the very small sample size of this dataset is a cause for concern, and these data should be taken with caution.

Additional stratified analyses based on age at onset and tumor characteristics (tables III and IV respectively), despite relatively small genotype strata for some SNPs, also shed light on susceptibility loci that may be useful for improved risk stratification and possible choice of therapy based on tumor characteristics. The contribution of rs7310409 to pNEN development is conceivable given that pleiotropic *HNF1A* plays a crucial role, as a tumor suppressor, in the transcriptional regulation in endocrine and exocrine pancreas development and homeostasis (31). This locus has also been reported as a prime candidate mediating susceptibility to type 2 diabetes among Caucasians (32). Although the role of p53 homolog, p63 in pancreatic cancer is poorly understood, it has been implicated in tumorigenesis and metastasis via cell cycle arrest and apoptotic mechanisms (33). Our observed association between heterozygous carriers of this polymorphism and metastatic pNEN tumors is therefore plausible.

Nevertheless, the small size observed in each stratum due to the rarity of these neoplasms, raises a valid concern on the possible interpretation of our findings. We therefore applied the FPRP approach instead of a much lower *P* value adjusted for multiple comparisons, as the latter would have resulted in unnecessarily low power for our high prior probability hypotheses, particularly in these neoplasms where collection of larger number of cases is unrealistic. Notwithstanding, all noteworthy findings reported here must be interpreted with caution and replicated in possibly larger and independent studies.

The major strengths of this study are its relatively large size, considering the low incidence of pNEN, and data analysis approaches applied to minimize the effect of potentially confounding factors and more importantly evaluate the noteworthiness of our findings. The integration of gene expression profiling provides valuable insight to shared genetic correlations and expands

the probable margins of our results. A limitation of our investigations however, was that due to lack of clinical information, we could not stratify subjects based on familial history of PDAC or other non-genetic factors such as history of diabetes mellitus.

Taken together, these results illuminate genetic similarities between the two main groups of pancreatic neoplasms, suggesting that common variations at three known PDAC susceptibility loci – 13q22.1 (*TP63*, rs9543325), 1q32.1 (*NR5A2*, rs10919791) and 8q24.21 (*MIR1208/PVT1*, rs1561927) potentially influence tumorigenesis of pancreatic beta cells in a similar manner as in exocrine pancreatic cells. Our results also suggest that genetic predisposing factors vary on the basis of age at pNEN onset as well as tumor characteristics. Striking similarities between embryogenesis and tumorigenesis which have become more apparent in past decades taken together with our genetic findings may provide groundwork for future evaluation of the potential usability of these loci as predictive biomarkers and in risk assessment. Functional studies are also warranted to better understand the biological implications of these loci in pancreatic hormone-producing (endocrine) and enzyme-producing (exocrine) cell function.

Acknowledgement

The authors acknowledge Angelika Stein (DKFZ, Heidelberg, Germany) for her technical support.

Author contributions

Conception and design: D.C, F.C and O.O.

Acquisition of samples and clinical data: All authors.

Analysis and interpretation of data (e.g., experimental and statistical analysis): O.O, D.C and F.C

Writing, review, and/or revision of the manuscript: O.O, G.C, D.C, F.C. and all other authors contributed to its critical revision.

References

1. Capurso G, Falconi M, Panzuto F, Rinzivillo M, Boninsegna L, Bettini R, et al. Risk factors for sporadic pancreatic endocrine tumors: a case-control study of prospectively evaluated patients. Am J Gastroenterol. 2009 Dec;104(12):3034–41.

2. Yao JC, Hassan M, Phan A, Dagohoy C, Leary C, Mares JE, et al. One Hundred Years After "Carcinoid": Epidemiology of and Prognostic Factors for Neuroendocrine Tumors in 35,825 Cases in the United States. J Clin Oncol. 2008 Jun 20;26(18):3063–72.

3. Li D, Xie K, Wolff R, Abbruzzese JL. Pancreatic cancer. Lancet Lond Engl. 2004 Mar 27;363(9414):1049–57.

4. Haugvik S-P, Hedenström P, Korsæth E, Valente R, Hayes A, Siuka D, et al. Diabetes, smoking, alcohol use, and family history of cancer as risk factors for pancreatic neuroendocrine tumors: a systematic review and meta-analysis. Neuroendocrinology. 2015;101(2):133–42.

5. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. Science. 2011 Mar 4;331(6021):1199–203.

6. Scarpa A, Chang DK, Nones K, Corbo V, Patch A-M, Bailey P, et al. Whole-genome landscape of pancreatic neuroendocrine tumours. Nature. 2017 Mar 2;543(7643):65–71.

7. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat Genet. 2009 Sep;41(9):986–90.

8. Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet. 2010 Mar;42(3):224–8.

9. Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet. 2014 Sep;46(9):994–1000.

10. Childs EJ, Mocci E, Campa D, Bracci PM, Gallinger S, Goggins M, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat Genet. 2015 Aug;47(8):911–6.

11. Campa D, Rizzato C, Stolzenberg-Solomon R, Pacetti P, Vodicka P, Cleary SP, et al. TERT gene harbors multiple variants associated with pancreatic cancer susceptibility. Int J Cancer. 2015 Nov 1;137(9):2175–83.

12. Campa D, Rizzato C, Capurso G, Giese N, Funel N, Greenhalf W, et al. Genetic susceptibility to pancreatic cancer and its functional characterisation: the PANcreatic Disease ReseArch (PANDoRA) consortium. Dig Liver Dis Off J Ital Soc Gastroenterol Ital Assoc Study Liver. 2013 Feb;45(2):95–9.

13. Bosman, Carneiro, Hruban, Theise. WHO Classification of Tumours of the Digestive System. 4th ed. World Health Organization ISBN: 978-92-832-2432-7; 2010. 417 p.

14. Wacholder S, Chanock S, Garcia-Closas M, Ghormli LE, Rothman N. Assessing the Probability That a Positive Report is False: An Approach for Molecular Epidemiology Studies. J Natl Cancer Inst. 2004 Mar 17;96(6):434–42.

15. Ramage JK, Ahmed A, Ardill J, Bax N, Breen DJ, Caplin ME, et al. Guidelines for the management of gastroenteropancreatic neuroendocrine (including carcinoid) tumours (NETs). Gut. 2012 Jan 1;61(1):6–32.

16. Klimstra DS, Modlin IR, Coppola D, Lloyd RV, Suster S. The pathologic classification of neuroendocrine tumors: a review of nomenclature, grading, and staging systems. Pancreas. 2010 Aug;39(6):707–12.

17. Park M, Kim M, Hwang D, Park M, Kim WK, Kim SK, et al. Characterization of gene expression and activated signaling pathways in solid-pseudopapillary neoplasm of pancreas. Mod Pathol Off J U S Can Acad Pathol Inc. 2014 Apr;27(4):580–93.

18. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 2013 Jan;41(Database issue):D991-995.

19. Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature. 2008 Oct 2;455(7213):627–32.

20. Hezel AF, Kimmelman AC, Stanger BZ, Bardeesy N, DePinho RA. Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev. 2006 May 15;20(10):1218–49.

21. Pour PM, Pandey KK, Batra SK. What is the origin of pancreatic adenocarcinoma? Mol Cancer. 2003 Jan 22;2:13.

22. Maitra A, Leach SD. Disputed paternity: the uncertain ancestry of pancreatic ductal neoplasia. Cancer Cell. 2012 Dec 11;22(6):701–3.

23. Deshpande V, Selig MK, Nielsen GP, Fernandez-del Castillo C, Lauwers GY. Ductulo-insular pancreatic endocrine neoplasms: clinicopathologic analysis of a unique subtype of pancreatic endocrine neoplasms. Am J Surg Pathol. 2003 Apr;27(4):461–8.

24. Landsman L, Nijagal A, Whitchurch TJ, Vanderlaan RL, Zimmer WE, Mackenzie TC, et al. Pancreatic mesenchyme regulates epithelial organogenesis throughout development. PLoS Biol. 2011 Sep;9(9):e1001143.

25. Regitnig P, Spuller E, Denk H. Insulinoma of the pancreas with insular-ductular differentiation in its liver metastasis--indication of a common stem-cell origin of the exocrine and endocrine components. Virchows Arch Int J Pathol. 2001 Jun;438(6):624–8.

26. Oliver-Krasinski JM, Kasner MT, Yang J, Crutchlow MF, Rustgi AK, Kaestner KH, et al. The diabetes gene Pdx1 regulates the transcriptional network of pancreatic endocrine progenitor cells in mice. J Clin Invest. 2009 Jul;119(7):1888–98.

27. McConnell BB, Yang VW. Mammalian Krüppel-Like Factors in Health and Diseases. Physiol Rev. 2010 Oct 1;90(4):1337–81.

28. Azmi A. Molecular Diagnostics and Treatment of Pancreatic Cancer: Systems and Network Biology Approaches. Elsevier; 2014. 467 p.

29. Robinson SR, Oliver AW, Chevassut TJ, Newbury SF. The 3' to 5' Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. Biomolecules. 2015 Jul 17;5(3):1515–39.

30. Kelly VR, Xu B, Kuick R, Koenig RJ, Hammer GD. Dax1 up-regulates Oct4 expression in mouse embryonic stem cells via LRH-1 and SRA. Mol Endocrinol Baltim Md. 2010 Dec;24(12):2281–91.

31. Hoskins JW, Jia J, Flandez M, Parikh H, Xiao W, Collins I, et al. Transcriptome analysis of pancreatic cancer reveals a tumor suppressor function for HNF1A. Carcinogenesis. 2014 Dec;35(12):2670–8.

32. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet. 2010 Jul;42(7):579–89.

33. Bergholz J, Xiao Z-X. Role of p63 in Development, Tumorigenesis and Cancer Progression. Cancer Microenviron Off J Int Cancer Microenviron Soc. 2012 Dec;5(3):311–22.

Table I. Clinical characteristics of cases and controls used in this study.

Region of origin	Germany	Greece	Italy	Poland	United Kingdom	Total subjects*
Median age at diagnosis†, *y* (1st-3rd quartile)	51 (44-58)	45 (35-58)	61 (49-72)	46 (31-66)	66 (59-74)	**57 (44-67)**
≤50 y	383	151	491	237	15	**1277**
>50 y	394	75	1238	185	209	**2101**
Gender						
Males	356	135	1098	185	130	**1906**
Females	421	95	814	257	94	**1681**
Functional status						
yes	4	20	47	3	3	**77**
no	n/a	2	118	3	n/a	**123**
unknown	31	0	82	9	47	**169**
Grade, G (WHO)						
G1 and G2	n/a	21	149	9	29	**208**
G3	n/a	1	6	2	1	**10**
unknown	35	0	12	4	20	**71**
Stage (ENETS)						
I / II / III	17	n/a	133	12	n/a	**162**
IV	5	n/a	32	2	n/a	**39**
unknown	13	22	82	1	50	**168**
Total cases	**35**	**22**	**247**	**15**	**50**	**369**
Total controls	**797**	**208**	**1670**	**427**	**175**	**3277**

Abbreviations: WHO = World Health Organization; ENETS = European Neuroendocrine Tumor Society; *y* = years; n/a = not available

* Numbers may not add up to 100% due to genotyping failure or unavailable covariate data.

† pNEN cases only

Table II. Associations between known GWAS PDAC susceptibility loci and pNEN risk.

SNP (nearby gene(s))	Genotypes (cases/controls[a])	MAF CEU 1KG / Current study[b]	MM vs Mm[c]		MM vs mm[c]	
			OR (95% CI)[d]	*P*	OR (95% CI)[d]	*P*
rs351365 (*WNT2B*)	CC (198/1549) CT (108/938) TT (26/172)	0.27/0.24	0.91 (0.69-1.20)	0.50	1.34 (0.82-2.18)	0.24
rs3790844 (*NR5A2*)	TT (140/1231) TC (64/721) CC (19/123)	0.36/0.23	0.82 (0.59-1.13)	0.22	1.27 (0.74-2.16)	0.38
rs1486134 (*ETAA1*)	TT (162/1569) TG (141/1120) GG (31/217)	0.27/0.27	1.20 (0.94-1.55)	0.13	1.36 (0.88-2.09)	0.16
rs9854771 (*TP63*)	GG (149/1131) GA (140/1311) AA (51/350)	0.29/0.36	0.76 (0.58-0.99)	0.04	1.00 (0.69-1.46)	0.98
rs2736098 (*TERT*)	GG (158/1514) GA (130/1131) AA (19/211)	0.27/0.27	1.08 (0.84-1.39)	0.56	0.84 (0.50-1.41)	0.51
rs17688601 (*SUGCT*)	AA (31/226) AC (130/1082) CC (176/1469)	0.17/0.28	0.98 (0.77-1.26)	0.89	1.19 (0.78-1.82)	0.41
rs6971499 (*LINC-PINT*)	AA (241/2222) AG (67/755) GG (5/72)	0.12/0.15	0.85 (0.64-1.14)	0.27	0.89 (0.35-2.27)	0.81
rs167020 (*SHH*)	AA (20/80) AG (108/392) GG (127/507)	0.15/0.26	1.04 (0.68-1.59)	0.85	0.92 (0.44-1.94)	0.83

rs1561927 (*MIR1208*, *PVT1*)	CC (25/238) CT (104/1297) TT (181/1643)	0.30/0.28	**0.71 (0.55-0.92)**	**0.01**	0.91 (0.58-1.43)	0.69

SNP (nearby gene(s))	**Genotypes (cases/controls[a])**	**MAF CEU 1KG / Current study[b]**	**MM vs Mm[c]**		**MM vs mm[c]**	
			OR (95% CI)[d]	***P***	**OR (95% CI)[d]**	***P***
rs10991043 (*SMC2*)	CC (48/371) CT (146/1295) TT (133/1190)	0.43/0.36	1.02 (0.79-1.31)	0.90	1.17 (0.81-1.68)	0.41
rs505922 (*ABO*)	TT (110/935) TC (120/1019) CC (28/333)	0.35/0.39	1.11 (0.84-1.48)	0.45	0.76 (0.48-1.19)	0.23
rs7310409 (*HNF1A*)	GG (120/1008) GA (159/1409) AA (59/490)	0.41/0.39	0.90 (0.69-1.16)	0.41	1.04 (0.74-1.46)	0.84
rs9581943 (*PDX1*)	GG (110/1016) GA (149/1524) AA (51/510)	0.33/0.42	0.94 (0.72-1.22)	0.64	0.91 (0.64-1.31)	0.62
rs9543325 *(KLF5/KLF12/PIBF1/DIS3/BORA)*	CC (48/316) CT (113/960) TT (62/771)	0.46/0.39	1.35 (0.96-1.89)	0.08	**1.63 (1.10-2.47)**	**0.02**
rs8028529 *(None)*	TT (155/1194) TC (86/654) CC (16/106)	0.20/0.22	1.08 (0.79-1.48)	0.64	1.25 (0.69-2.25)	0.46
rs7190458 (*BCAR1/CTRB1/CTRB2*)	CC (294/2832) CT (18/215) TT (1/5)	0.10/0.04	0.78 (0.46-1.30)	0.34	2.25 (0.25-20.3)	0.47

rs11655237 (*LINC00673*)	CC (256/2113) CT (75/583) TT (3/37)	0.23/0.12	1.08 (0.80-1.45)	0.63	0.70 (0.16-3.02)	0.64
rs1517037 (*GRP*)	CC (227/1871) CT (107/896) TT (10/129)	0.24/0.20	0.89 (0.69-1.15)	0.36	0.59 (0.30-1.15)	0.12
rs16986825 (*ZNRF3*)	CC (193/1996) CT (100/919) TT (15/108)	0.20/0.19	1.07 (0.82-1.39)	0.63	1.21 (0.67-2.17)	0.53
rs2853677 (*TERT*)	AA (89/783) AG (155/1319) GG (52/532)	0.39/0.45	1.01 (0.74-1.38)	0.94	0.82 (0.54-1.22)	0.32
rs401681 (*TERT*)	TT (51/377) TC (129/949) CC (76/649)	0.41/0.43	1.10 (0.81-1.52)	0.52	1.11 (0.75-1.63)	0.61
rs10919791 (*NR5A2*)	GG (148/1272) GA (50/682) AA (15/111)	0.35/0.22	**0.65 (0.46-0.93)**	**0.02**	1.06 (0.59-1.90)	0.84
rs3790843 (*NR5A2*)	GG (124/1064) GA (72/847) AA (25/178)	0.40/0.29	0.76 (0.55-1.04)	0.09	1.20 (0.74-1.93)	0.46

SNP (nearby gene(s))	**Genotypes (cases/controls[a])**	**MAF CEU 1KG / Current study[b]**	**MM vs Mm[c]**		**MM vs mm[c]**	
			OR (95% CI)[d]	***P***	**OR (95% CI)[d]**	***P***

Abbreviations: MAF CEU = minor allele frequency in Caucasians; 1kG = 1000 Genomes Project (http://browser.1000genomes.org); OR = odds ratio; CI = confidence interval.

Only association results for SNPs with P<0.05 and with outcomes below FPRP-level 0.2 are marked bold.

[a] Numbers may not add up to 100% due to genotyping failure or unavailable covariate data; [b] Controls only; [c] M - major allele, m - minor allele where MM represents the reference category.

[d] Analyses were performed with adjustment for age, gender and country of origin.

Table III. Associations between *TERT*_rs2736098, *ZNF*_rs16986825, rs9543325 and risk of pNEN onset.

Locus	Gene	SNP ID	Genotypes (Cases≤50y/>50y)	OR (95% CI)[a]	*P*
5p15.33	*TERT*	rs2736098	GG (52/104)	0.83 (0.50-1.40)†	0.49
			GA (36/89)	**3.71 (1.36-10.10)***	**0.01**
			AA (12/7)	1.04 (0.64-1.69)¥	0.89
22q12.1	*ZNRF3*	rs16986825	CC (53/136)	**2.67 (1.57-4.55)†**	**0.0003**
			CT (46/52)	0.67 (0.17-2.63)*	0.57
			TT (3/11)	**2.27 (1.37-3.77)¥**	**0.002**
3q28	*TP63*	rs9854771	GG (38/108)	**1.85 (1.11-3.10)†**	**0.02**
			GA (53/87)	1.29 (0.63-2.66)*	0.49
			AA (15/35)	1.69 (1.04-2.75)¥	0.03

Only association results for SNPs with P<0.05 and with outcomes below FPRP-level 0.2 are marked bold.

[a] Cases only - ≤50 years (n=114) versus >50 years (n=247); where cases >50 years represent the reference category. Analyses were performed with adjustment for gender and country of origin.

† MM vs Mm; * MM vs mm; ¥ (MM+Mm) vs mm; where M - major allele and m - minor allele; y= years.

Table IV. Associations between SNPs and pNEN tumor characteristics.

Locus	Gene	SNP	Genotype groups	OR (95% CI)[a]	*P*
Grade (G1/G2/well differentiated tumors vs. G3/poorly differentiated tumors)					
1q32.1	*NR5A2*	rs10919791	GG/GA/AA 115/38/15	**5.66 (1.18-27.11)**†	**0.03**
22q12.1	*ZNRF3*	rs16986825	CC/CT/TT 113/62/9	7.34 (1.02-52.67)* 7.44 (1.10-50.53)¥	0.05 0.04
Stage (I/II/III vs. IV)					
8q24.21	*MIR1208*, *PVT1*	rs1561927	CC/TC/TT 18/55/95	**2.39 (1.18-4.85)**†	**0.02**
3q28	*TP63*	rs9854771	AA/GA/GG 22/67/81	**0.44 (0.22-0.90)**†	**0.02**
1p13.1	*WNT2B*	rs351365	AA/GA/GG 11/58/99	4.15 (1.01-16.95)* **4.62 (1.16-18.38)**¥	0.05 **0.03**
Functional status (functional tumors vs. non-functional tumors)					
12q24.21	*HNF1A*	rs7310409	AA/GA/GG 32/79/66	**0.36 (0.16-0.79)**†	**0.01**
18q21.2	*GRP*	rs1517037	CC/CT/TT 118/56/4	**0.37 (0.16-0.85)**†	**0.02**

Only association results for SNPs with P<0.05 and with outcomes below FPRP-level 0.2 are marked bold. Analyses were performed with adjustment for age, gender and country of origin.

[a] All G1 and G2 (i.e well-differentiated) tumors, non-functional tumors, and Stage I, II and III tumors (i.e. no nodal metastasis) represent the reference category for *Grade, Functional status* and *Stage* respectively. † MM vs Mm; * MM vs mm; ¥ (MM+Mm) vs mm; where M - major allele and m - minor allele

Table V. Differentially-expressed genes obtained from the GSE43795 dataset (6 pNEN tumors vs. 5 non-neoplastic pancreatic tissue samples and 6 pNEN tumors vs. 6 ductal adenocarcinoma tissues).

Gene	pNEN vs. normal			PDAC vs. normal		
	Log2 fold-change	*P*	Adjusted-*P**	Log2 fold-change	*P*	Adjusted-*P**
MYC	4.38	9.47×10^{-7}	5.88×10^{-5}	2.65	4.55×10^{-4}	5.93×10^{-3}
NR5A2	4.91	7.82×10^{-10}	2.38×10^{-7}	3.51	2.00×10^{-4}	3.31×10^{-3}
KLF5	1.69	1.59×10^{-2}	6.96×10^{-2}	-1.41	1.59×10^{-1}	3.18×10^{-1}
CTRB1	7.90	1.10×10^{-9}	3.04×10^{-7}	7.55	2.56×10^{-8}	5.18×10^{-6}
CTRB2	7.70	6.75×10^{-8}	8.08×10^{-6}	7.33	1.71×10^{-6}	1.13×10^{-4}
PDX1	0.72	3.58×10^{-1}	5.49×10^{-1}	2.43	1.18×10^{-3}	1.16×10^{-2}

* *P*-value adjusted using Benjamini & Hochberg method (False discovery rate)

Based on predefined cutoff value of Adjusted-*P* < 0.05 and |log2 fold-change| > 1, genes which are differentially expressed in both groups are in bold.